From Scientist to Data Scientist

Joseph Pellegrino, PhD Quantum Optics
Chief Data Scientist @ Kernix

From Science to Data Science, May 2017

The Arrow of Time

2011 2012 2013 2014 2015 2016 2017

Phase transition ?

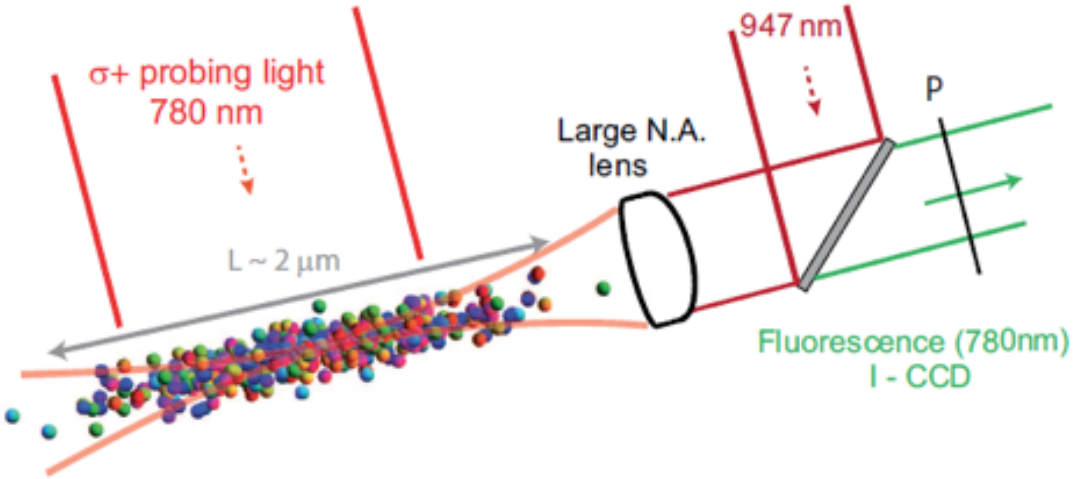
PhD candidate in Quantum Physics
@ LCF of IOGS

Data Scientist @ Kernix

From Scientist ...

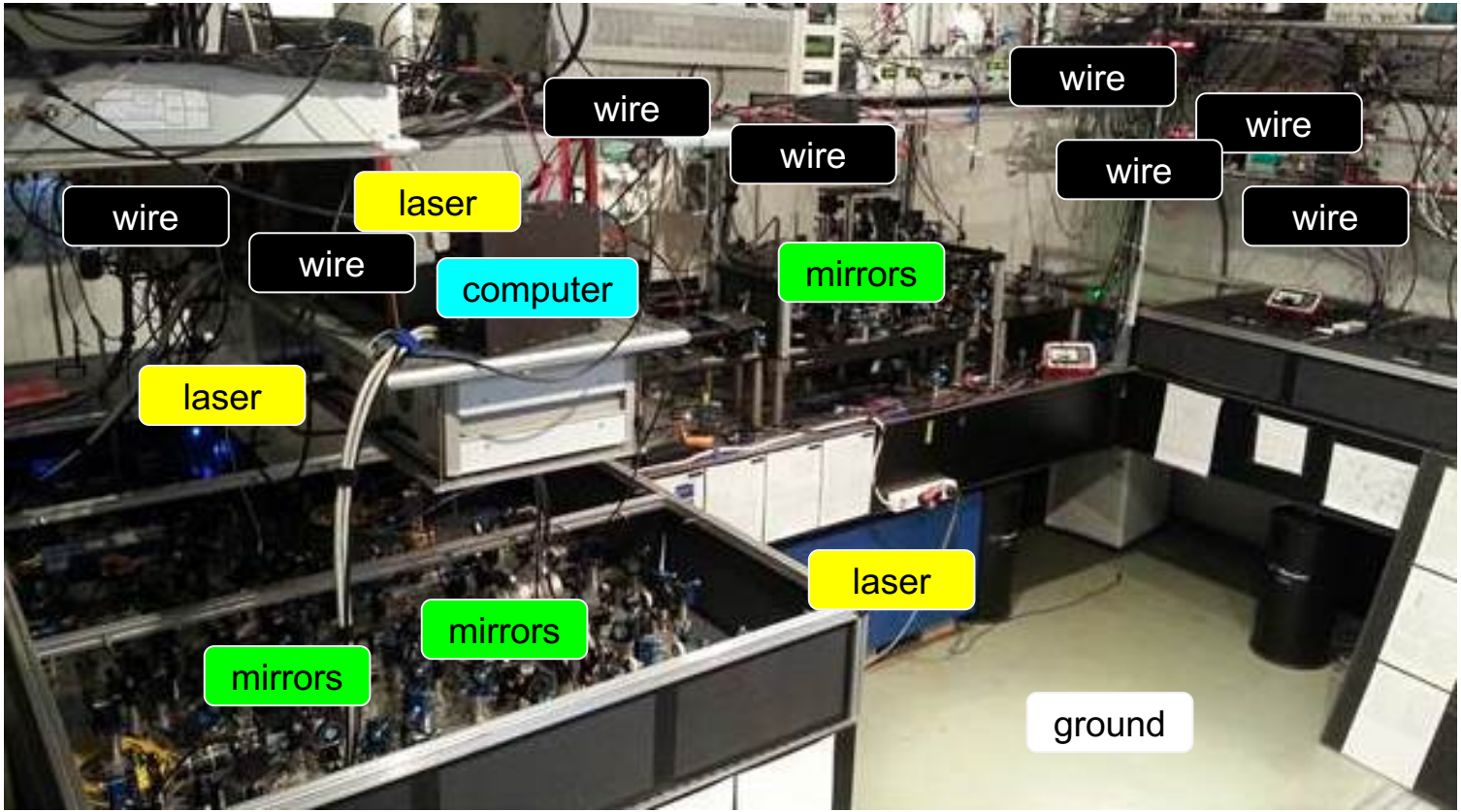
A PhD thesis on ...

How does small and dense clouds of cold atoms scatter resonant light ?



I was looking for *collective* behavior induced by *interactions between entities*...

Welcome in the cold atom lab !



Some scientific skills

Experimental	*Theoretical*
Optics (detectors, lasers, electro & acousto-optic modulator, ...)	Classical and quantum optics
Electronics (control systems, monitoring systems, …)	Atomic physics
Vacuum engineering (ultra-high vacuum chamber, ionic pumps,..)	Statistical physics
Computer system (synchronisation of experiments, data acquisition & analysis, …)	Signal processing

October 2014 : PhD defense

Phase transition ?

What I love :

Understanding things

Physics & Mathematics

Solving hard problems

Building LEGO systems

Programming

Helping people to do great stuff

What I would like to do:

What I love

Something in phase with our everyday life

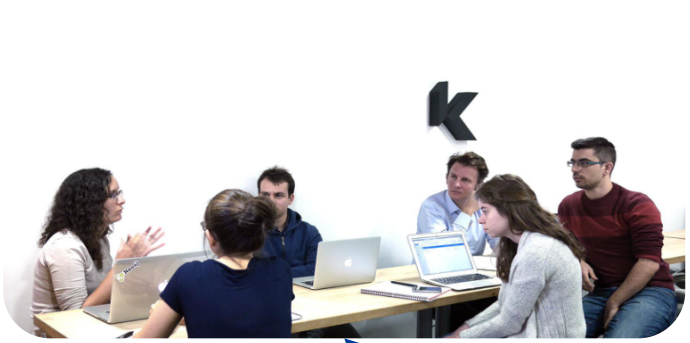
Something quickly useful for people

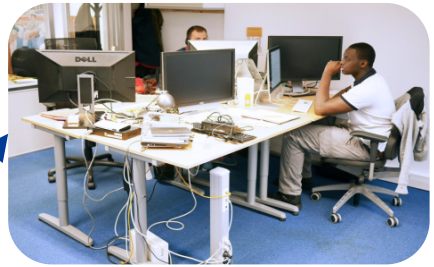
And one day, I heard about the “Big Data” trend ...

Looking for data science positions ...

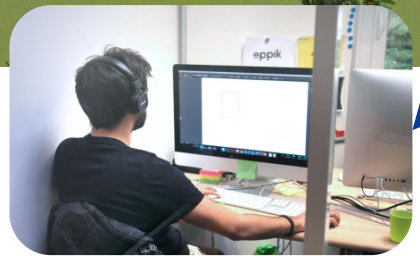
... to Data Scientist.

Kernix: Digital Factory and Data Lab





I was looking for *collective* behavior induced by *interactions* between entities...



DATA LAB

Clients

Schneider Electric · PTC · SNCF RÉSEAU · KONBINI · Cdiscount.com · PRICEMINISTER · Performics

Collaborations

EPFL ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE · HEC PARIS · EDC PARIS BUSINESS SCHOOL · MINES ParisTech · cap-digital · eit Digital

EXPERTISE

Data Pipelines

- Cop21
- TerraRush

Predictive maintenance

- ERDF

Data Vizualisation

- SolarImpulse

Recommender systems

- PriceMinister
- WikiDistrict
- Clickalto
- HobbyStreet

Marketing Automation

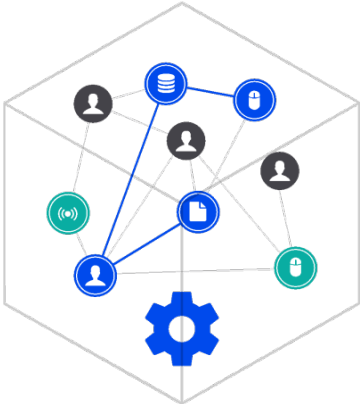
- Performics
- RadiumOne

Open Data

- Accessible.net

Steps of a Data project

DEFINE → COLLECT → ANALYSE → EXPLOIT



New skills to develop

Hard skills :

- Programming languages (Python, Javascript)
- Programming style (Object Oriented, Functional)
- Databases (MongoDB, Neo4j, MySQL, Redis)
- Machine learning
- Image processing
- Natural Language Processing
- Graph mining

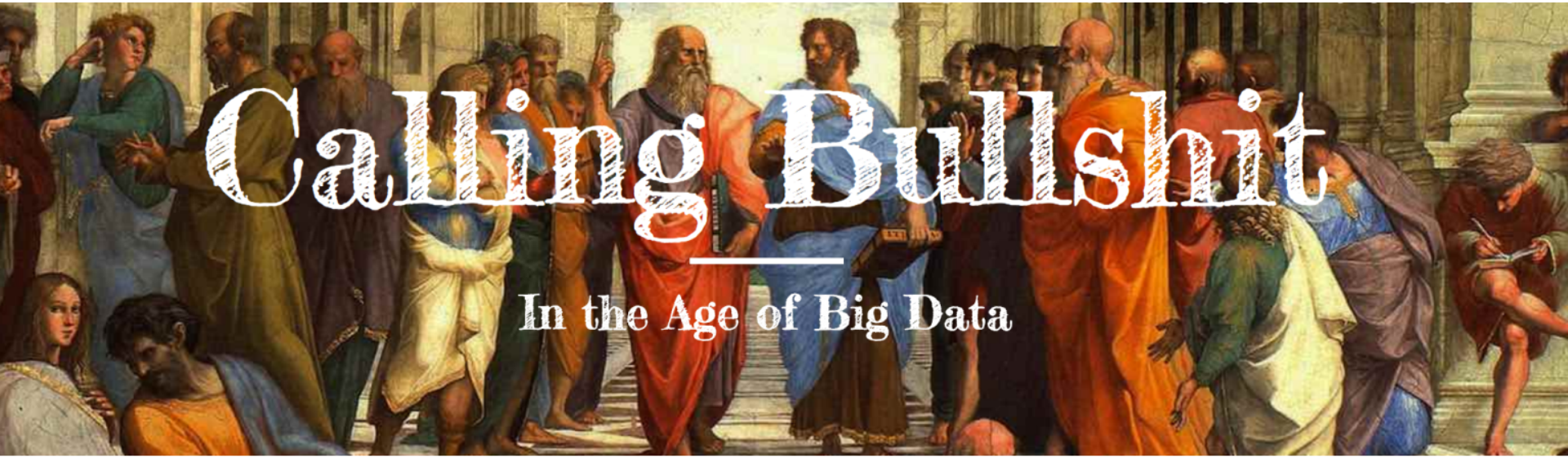
Soft skills :

- Project management
- Business understanding
- Team management

A shift of paradigm

Science

Data Science



Critical and analytical thinking : in order to eradicate bullshit

Are Data Sciences an ephemeral hype ?

Thank you for your attention